



CAMBRIDGE
UNIVERSITY PRESS

Political Confederation

Author(s): Jacques Crémer and Thomas R. Palfrey

Source: *The American Political Science Review*, Vol. 93, No. 1 (Mar., 1999), pp. 69-83

Published by: [American Political Science Association](#)

Stable URL: <http://www.jstor.org/stable/2585761>

Accessed: 08-03-2016 00:52 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Political Science Association and Cambridge University Press are collaborating with JSTOR to digitize, preserve and extend access to *The American Political Science Review*.

<http://www.jstor.org>

Political Confederation

JACQUES CRÉMER *Université des Sciences Sociales de Toulouse*

THOMAS R. PALFREY *California Institute of Technology*

This article extends the spatial model of voting to study the implications of different institutional structures of federalism along two dimensions: degree of centralization and mode of representation. The representation dimension varies the weight between unit representation (one state, one vote) and population-proportional representation (one person, one vote). Voters have incomplete information and can reduce policy risk by increasing the degree of centralization or increasing the weight on unit representation. We derive induced preferences over the degree of centralization and the relative weights of the two modes of representation, and we study the properties of majority rule voting over these two basic dimensions of federalism. Moderates prefer more centralization than extremists, and voters in large states generally have different preferences from voters in small states. This implies two main axes of conflict in decisions concerning political confederation: moderates versus extremists and large versus small states.

Episodes of political confederation or the opposite, political dissolution, are widely observed and important phenomena. Yet, the forces underlying the processes of confederation are not well understood. There are many different levels of political units, ranging from towns and villages to states or provinces, to nation-states, to supranational bodies such as the United Nations and the European Union. Generally rules exist that govern both the interaction between units of different levels and the allocation of decision-making authority among them. It is the determination of such rules and the development of a theoretical foundation for studying them that we address in this article. We call this the problem of *political confederation*.

Political confederation is one of the most fundamental problems in political science. It arises in many contexts, both large and small. Several examples come immediately to mind:

- the ceding of sovereignty to international organizations by member nations;
- the issue of "states rights" in American politics;
- the nature of representation of provinces in national assemblies;
- the allocation of taxing authority in a federal system;
- the adoption of uniform standards across jurisdictions (pollution, education, roads, etc.);

- the formation of nation-states; and
- the rules for amending constitutions.

A number of specific examples have been studied in great detail by scholars of history and political science, particularly with regard to the evolution and collapse of international organizations and the drafting of constitutions. Most of this work attempts to describe the important forces that were operative in those isolated incidents, and it provides some useful insights into the problem. Here we attempt to capture some general features of the problem in a formal model and to derive some theoretical implications.

We model a confederation as a collection of districts, together with rules for the aggregation of preferences of the members of the various districts, which produces policy outcomes that can vary across districts. The districts can be thought of as regions, provinces, counties, villages, or other well-defined decentralized political units. The larger body is referred to as the "confederation," and it can be thought of as any political unit that contains some collection of smaller political units.

In terms more familiar to economists than political scientists, it is a mechanism design problem, that is, a metadecision-making problem in which the decision to be taken is a set of rules for making decisions. It can also be thought of as the problem of constitutional design. What makes the problem of constitutional design especially complicated is that the designers themselves have preferences over the outcomes and policies that ultimately will be decided according to the rules for preference aggregation that are chosen at the constitution-making stage.¹ Thus, the designers will have *induced preferences* over the rules that will depend on their preferences over policy outcomes and the specific features of their own district.

Jacques Crémer is Directeur de Recherche au CNRS, GREMAQ-IDEI, Université des Sciences Sociales de Toulouse, 31042 Toulouse Cedex, France. Thomas R. Palfrey is Professor of Economics and Political Science, California Institute of Technology, Pasadena, CA 91125.

This research was supported in part by the National Science Foundation through grants SES-9224787 and SBR-9631627. The work was begun while Palfrey was visiting Université de Toulouse in 1995. Its financial support, as well as its stimulating research environment, is gratefully acknowledged. Palfrey also thanks CREST-LEI and CERAS for their hospitality and research support during the 1995–96 academic year.

We are grateful for comments from R. Michael Alvarez, Jonathan Katz, and seminar participants at London School of Economics, Université de Toulouse, the 1996 Gerzensee Workshop on Political Economy, the 1997 SITE summer program on Interregional Competition in Public Economics, the 1997 annual meeting of the Public Choice Society, and the referees.

¹ See the historical analyses of the formation of the U.S. Constitution, such as Riker (1984, 1986). Grofman and Wittman (1989) present a collection of articles that emphasize the important connection between policy preferences and induced preferences over institutions. These and other studies provide convincing evidence that this link played a prominent role in the design of the U.S. Constitution.

Our simple framework combines the standard spatial model of political competition with the presence of asymmetric information and multiple districts. It builds on the model of Cr mer and Palfrey (1996), in which voters from a district know only their own ideal point in a policy space and some common information about the distribution of preferences in their district and in other districts. With incomplete information, the policy outcomes that emerge from different rules of confederation are uncertain prospects, so the induced preferences over confederations reflect not only a voter's ideal point but also a voter's attitudes about risky outcomes. In the standard Downsian model, preferences necessarily exhibit some degree of concavity with respect to outcomes, which produces risk aversion if the policy outcomes are uncertain.

Cr mer and Palfrey (1996) considered only two polar cases. At one extreme, the null confederation, there is no national government at all, and the policy outcomes of different jurisdictions are determined independently, simply as the median preference of the district. At the opposite extreme, the districts cede all sovereignty to a central authority, so that all districts have the same policy, which is decided by the median of the nation (composed of all the districts, aggregated together). Because voters do not know the preferences of other voters, the two policy outcomes under the different systems (independent versus confederated) are random variables that differ in both their conditional mean and variance. Because preferences are quadratic and the uncertainty is normally distributed, the induced preferences depend only on the mean and variance, and this leads to clean analytical solutions.

The basic trade-off for voters is that centralization leads to policy outcomes that are, on average, farther from their ideal point, while the risk associated with the centralized government is lower than the risk associated with independent policymaking. The latter follows from the fact that the component of risk due to interdistrict variance of policy outcomes is reduced with confederation. We call this risk reduction the *principle of moderation*. As far as we know, this principle has never been formalized before, but it is not new, for it played a central role in Madison's arguments for the unification of the American states in the late eighteenth century. For example, it is related to the problem of extreme political factions discussed in *Federalist #10* (Madison et al. [1788] 1937, 61–2):

The influence of faction leaders may kindle a flame within their particular States, but will be unable to spread a general conflagration through the other States. A religious sect may degenerate into a political faction in a part of the Confederacy; but the variety of sects dispersed over the entire face of it must secure the national councils against any danger from that source. A rage for paper money, for an abolition of debts, for an equal division of property, or for any other improper or wicked project, will be less apt to pervade the whole body of the Union than a particular member of it; in the same proportion as such a malady is more likely to taint a particular county or district, than an entire State.

The central finding in Cr mer and Palfrey (1996) is that moderate voters prefer centralization and extreme voters oppose it. Specifically, there is a critical voter ideal point such that all voters whose ideal point is closer to the expected median prefer centralization, and all voters whose ideal point is farther away from the expected median prefer separatism. We extend this model in two ways.

First, we consider a much broader range of constitutions than the two extreme cases of “all or nothing” national government. These intermediate cases are obtained by representing the policy outcome in a district along a continuum of possible mixes between purely federal and purely local policy. We call this the *centralization dimension* and show below that it captures a wide range of “federal” systems, in which the relative power of a district in affecting policies in other districts is allowed to vary.

Second, we introduce another dimension to the confederation design, the *representation dimension*. We consider two contrasting notions of representational formulas, one in which a local unit's representation at the national level is proportional to the population of the unit, termed *population-proportional representation*,² and one in which each district receives the same absolute representation, termed *unit representation*. An example of these two extremes is found in the U.S. Congress; if we think of states as the basic units of the confederation, the House of Representatives approximates population representation, and the Senate has unit representation. Hereafter, we sometimes refer to these two types of representation as “House” and “Senate,” respectively. Our representation dimension considers linear combinations of these two schemes, which can loosely be thought of as the allocation of power between two different chambers in a bicameral national legislature (again using the U.S. analogy).

Madison also discusses the trade-off between unit and population representation. While the latter provides greater retention of local sovereignty for large states, the former is a more effective way to moderate factions, especially those from large states. Madison addresses the issue of why one ought to have unit (Senate) in addition to population (House) representation in the legislative body,³ and he realizes that unit representation of states helps resolve a trade-off between loss of sovereignty and moderation, in a way that all states (large and small) should come to recognize it as a good solution to the representation problem. Here is an example of Madison's discussion, taken from *Federalist #62* (Madison et al. [1788] 1937, 401–2).⁴

² For the rest of the article, we shorten this to *population representation*.

³ His terms are “equal” and “proportional” representation, respectively. Because modern usage of those terms has changed, we adopt the terms unit representation and population representation to avoid confusion.

⁴ Madison was also well aware that small states would be more likely to ratify a constitution that provided for unit as well as population representation: “The equality of representation in the Senate is another point, which, being evidently the result of compromise between the opposite pretensions of the large and the small States, does not call for much discussion A government founded on

...in a compound republic, partaking both of the national and federal character, the government ought to be founded on a mixture of the principles of proportional [population] and equal [unit] representation... In this spirit, it may be remarked, that the equal vote allowed to each State is at once a constitutional recognition of the portion of sovereignty remaining in the individual states, and an instrument for preserving that residuary sovereignty.

Accordingly, we examine these various trade-offs in a formal theoretical model by characterizing induced voter preferences over a two-dimensional set of possible constitutions, with centralization and representation as the two axes. In the context of a normal-quadratic model,⁵ we obtain the following results.

1. Voter preferences over representation depend on the degree of centralization, and vice versa. As the level of centralization increases, voters from large states prefer more population representation (i.e., more House power), while voters from small states prefer less population representation (i.e., more Senate power). As the weight on population representation increases, voters from large states prefer more centralization, while voters from small states prefer less centralization.
2. Each voter has an induced ideal point that consists entirely of unit representation with a most-preferred level of centralization that depends on the voter's ideal point in the underlying policy space but not on the district size. Interestingly, this applies equally to voters from large states (relative population size greater than $1/n$), who sacrifice power to small states under unit representation.
3. Consistent with the findings of Cr mer and Palfrey (1996), more extreme voters want less centralization.
4. A majority rule equilibrium over the two-dimensional confederation issue exists within each district.
5. An interdistrict majority rule equilibrium over the two-dimensional confederation issue always exists in the case of two districts. The equilibrium mix of representation can range from completely population based to completely unit based, including all mixtures in between. The exact mixture depends on the underlying distributions of voter preferences and of district sizes.
6. Conditional on any level of centralization, preferences are single peaked in the representation dimension, for any number of districts. Thus, there always exists a conditional majority rule equilibrium, which typically involves a mix of population and unit representation.
7. An equilibrium across all voters in all districts may or may not exist with more than two districts. Several examples are given to illustrate the range of possibilities.

principles more consonant to the wishes of the larger States, is not likely to be obtained from the smaller States" (Madison et al. [1788] 1937, 401). This property is also captured in our model.

⁵ We discuss below the robustness of our results to alternative specifications.

Result 1 is quite intuitive. When voters think there is too much power in the central government, they wish to have greater influence on central decisions. Voters in small states have more say under a system of unit representation than population representation, whereas the reverse is true for voters in large states. Results 2 and 3 follow from the basic intuition developed in Cr mer and Palfrey (1996). All voters benefit from centralization due to its mitigating effect on policy risk.⁶ This risk is reduced more efficiently (i.e., with a less powerful central government) if representation is unit based, simply because equal shares maximize the diversification.⁷ Result 4 follows from the fact that the indifference curves of voters from the same district are lined up symmetrically. Result 5 follows from the fact that with two districts voter indifference curves are straight lines, so the problem is virtually one dimensional. Result 6 is true because indifference curves are convex, with unique conditional ideal points. The non-existence possibility in result 7 derives from two properties of the model: (1) the strong interaction of individual preferences between the centralization and representation dimension, and (2) the fact that this interaction works in opposite ways for small and large districts.

Results 5 and 6 are particularly noteworthy because they show that our theoretical model can account for the variety of balances between population and unit representation observed at the national level. Moreover, our model gives some indication of how this mix varies with the underlying distribution of preferences, the number of subnational regions, and the size distribution of regions. This is discussed and documented in Tsebelis and Money (1997), in the context of their comparative study of bicameral legislative systems. One striking feature, for example, is that it is common for upper houses in bicameral systems to have unit representation: "In upper houses where local political units (states, cantons, regions, departments, or counties) are granted unit representation, such as the United States and Switzerland, some citizens' votes are weighted much more heavily than others. This pattern is visible in approximately one-third of all bicameral legislatures" (Tsebelis and Money 1997, 53). This is also observed in supranational organizations such as the European Parliament and the United Nations. We know of no other formal model that accounts for the ubiquity of unit representation and at the same time is consistent with the empirical fact that mixtures of unit and population representation are commonly observed.

The rest of the article is organized as follows. In the next section, we introduce the model and notation. We then characterize the induced preferences over the two-dimensional space of confederations—representation and centralization. Finally, we present our results

⁶ This statement applies to both extreme and moderate voters. For sufficiently extreme voters, however, this risk-reduction benefit is outweighed by the fact that the expected centralized outcome is very far from their ideal point.

⁷ If districts had different variances, then equal shares might not lead to the most efficient diversification of risk.

on the existence and properties of majority rule equilibria. Formal proofs are given in the Appendix.

THE MODEL

A *confederation* is composed of n districts, in each of which there is a continuum of individuals. The relative size of district i is denoted $\alpha_i > 0$, where $\sum_{i=1}^n \alpha_i = 1$. The underlying policy space is one dimensional and represented by \mathfrak{R} , the set of real numbers. We assume that the political process within any district yields a policy that is the ideal point of the median voter of that district.

Each individual has an *ideal policy*, $t \in \mathfrak{R}$, which we also call his *type*. If policy $x \in \mathfrak{R}$ is adopted, the utility of an individual of type t is

$$U(x; t) = -(x - t)^2.$$

When voting over confederations, the agent is uncertain about the outcome. In this case, between two distributions over policies he prefers the distribution with the higher expected utility.

An individual is represented by his type, t , and the district, i , to which he belongs. The distribution of voter preferences is determined by two components. The first is the vector of district medians, each of which is an independent draw from a normal distribution, with mean 0 and variance V . The vector of district medians is denoted $m = (m_1, \dots, m_n)$. Within district i , the ideal policies (types) of the continuum of voters follow a normal distribution, with mean m_i and variance equal to 1. Thus, there is a partial sorting of preferences of voters, by district, since the ideal points of district i are centered around that district's median, and (with probability one) this median is different from the median of each of the other $n - 1$ districts. Each individual only knows her own type, the district to which she belongs, the relative size of all districts, and V , but she does not know the vector of district medians.⁸

Confederations

A confederation is an institutional arrangement in which the policies of different districts are, at least in part, influenced by the preferences of voters from other districts in the confederation. In practice, this is usually accomplished through a complex array of overlapping jurisdictions, representative governments at different levels, and a legal system that allocates decision-making authority and responsibility across these different levels. The end result is ultimately a vector of policies, one for each district. We denote such a vector of policies by $x = (x_1, \dots, x_n)$.

Given the set of institutions and the legal system

governing the overlapping jurisdictions, we can think of x as emerging as a function of the underlying preferences in all the different districts. Since we take "district" to be the smallest political unit, and since we assume that political competition within a district will be driven by the median voter of that district, we write

$$x = C(m_1, \dots, m_n) \\ = [C_1(m_1, \dots, m_n), \dots, C_n(m_1, \dots, m_n)].$$

That is, the constitution of a confederation is modeled as a function C , which maps vectors of district medians into vectors of district policies. This function may vary among districts and generally will be different in the applications we explore below.

To make the analysis tractable, we limit consideration to confederations in which $C_i(m_1, \dots, m_n)$ is linear:

$$C_i(m_1, \dots, m_n; \lambda) = \sum_{j=1}^n \lambda_{ij} m_j,$$

with $\lambda_{ij} \geq 0$, and $\sum_j \lambda_{ij} = 1$. That is, we make the simplifying assumption that a confederation can be represented by a matrix of *influence coefficients*, $\lambda = [\lambda_{ij}]$, where λ_{ij} is the influence of district j on the policy outcome in district i .

The Representation/Centralization Axes

Using this notation, the two special cases considered in Cr mer and Palfrey (1996) were the following. At one extreme was $C_i(m_1, \dots, m_n) = m_i$ (independent districts), and at the other extreme was a size-weighted average of the district means:⁹

$$C_i(m_1, \dots, m_n) = \sum_{j=1}^n \alpha_j m_j, \text{ for all } i.$$

That article analyzed the induced individual preferences between these two extreme confederations and identified conditions under which a majority of individuals (or a majority of districts) would prefer unification to independence.¹⁰ Here, we study the induced prefer-

⁹ Since we have a normal distribution of voter ideal points within a district, the median ideal point in a district coincides with the mean ideal point in that district. Modelling the unified policy as a weighted mean simplifies the analysis and captures the main idea that the centralized policy will represent a compromise between the medians of the individual districts.

We do not model the details of the policymaking process but note that the literature identifies many different circumstances in which outcomes are responsive to the mean rather than the median and/or in which outcomes vary continuously with the preferences of the members of the legislature. Examples include most models with uncertainty or private information (Ledyard 1984), probabilistic voting (Coughlin 1990; Hinich 1977), divided government (Alesina and Rosenthal 1995), and distributive politics (Banks and Duggan 1998; Baron and Ferejohn 1989; Merlo 1997).

¹⁰ Using the notation of influence coefficients, these two extremes correspond to $\lambda_{ii} = 1$ for all i (independence) and $\lambda_{ij} = \alpha_j$ for all i and j (unification).

⁸ This can be generalized in a number of ways. For example, voters could have additional information about their own district, there could be correlation between m_i and α_i , or there could be different within-district variances. See Cr mer and Palfrey (1996) for a discussion of these and other possible extensions.

ences of individuals over a much richer set of possible confederations.

In this article, we define unification more generally by $\lambda_{ij} = \lambda_j$ for all i and j , with λ_j not necessarily equal to α_j . The case of $\lambda_{ij} = \alpha_j$, treated in our earlier article, corresponds to *population representation*. *Unit representation* is defined by $\lambda_{ij} = 1/n$ for all i, j . We can represent a continuum of multicameral unified systems using a parameter, $\gamma \in [0, 1]$, that indexes the relative influence of population representation. In a unified system this would correspond to $\lambda_{ij} = \gamma\alpha_j + (1 - \gamma)/n$. This defines what we call the *representation dimension* and spans all ranges of mixtures between purely population representation ($\gamma = 1$) and purely unit representation ($\gamma = 0$).

Of course, nearly all unified confederations of states also have governments at the district level, so they are a cross between the two extremes of pure independence and full unification. We represent a continuum of degrees of centralization using a parameter $\beta \in [0, 1]$, which represents the relative weight of the central government in determining policy outcomes in any given district. This defines what we call the *centralization dimension*, and spans all ranges of mixtures between purely independent systems ($\beta = 0$) with no central government and fully unified systems ($\beta = 1$) in which districts have no autonomy at all.

Putting these two dimensions together, we have the two main axes—representation and centralization—that determine our space of confederations. In the notation of influence coefficients, the policies in each district are determined according to

$$\lambda_{ii} = 1 - \beta + \beta(\gamma\alpha_i + (1 - \gamma)/n),$$

$$\lambda_{ij} = \beta(\gamma\alpha_j + (1 - \gamma)/n).$$

This restricts the set of admissible matrices of influence coefficients. Specifically, $\alpha_i = \alpha_{i'}$ implies $\lambda_{ij} = \lambda_{i'j}$ for all i, i' , and j that are not all equal to one another. In other words, the influence of a district on the policy outcome in another district is independent of the identity (but not the size) of that second district.

The actual policy outcome in district i , $C_i(m_1, \dots, m_n; \alpha, \beta, \gamma)$, is given by

$$\begin{aligned} C_i(m_1, \dots, m_n; \alpha, \beta, \gamma) &= \sum_{j \neq i} m_j \beta(\gamma\alpha_j + (1 - \gamma)/n) \\ &\quad + m_i[1 - \beta + \beta(\gamma\alpha_i + (1 - \gamma)/n)] \\ &= (1 - \beta)m_i + \beta \left[\gamma \sum_{j=1}^n \alpha_j m_j + (1 - \gamma) \sum_{j=1}^n m_j/n \right] \\ &= (1 - \beta)m_i + \beta[\gamma M_h + (1 - \gamma)M_s], \end{aligned}$$

where $M_h = \sum_{j=1}^n \alpha_j m_j$, and $M_s = \sum_{j=1}^n m_j/n$ can be interpreted as the *House policy* and the *Senate policy*, respectively.

The rest of the article will characterize induced voter

preferences and majority rule equilibria in this two-dimensional space of confederations.

INDUCED PREFERENCES OF VOTERS

In order to understand political conflict along these two basic dimensions of the confederation question, in this section we characterize the induced preferences of voters over the (β, γ) space, represented by the square $[0, 1] \times [0, 1]$. This characterization identifies the shape of voter indifference curves in this space and shows how the ideal points of voters in this space vary as a function of n, i, V, t , and the vector α .

A voter is characterized both by his ideal point, t , and his district, i , and we will speak of the voter (i, t) . To simplify the formulas and the exposition, we define $\theta = t^2/(1 + V)$, which is a measure of how extreme a voter's ideal point is relative to the unconditional expected ideal point. We use the term "voter (i, θ) " to refer to the two voters, $(i, \sqrt{\theta(1 + V)})$ and $(i, -\sqrt{\theta(1 + V)})$, who are equally extreme and hence (as we formally prove below) have identical induced preferences over confederations. For such a voter, the utility of confederation (β, γ) equals $-W_{i\theta}(\beta, \gamma)$, where $W_{i\theta}(\beta, \gamma)$ is the expected squared distance between voter (i, θ) 's ideal policy and the outcome under (β, γ) . Through standard computations similar to those detailed in Cr mer and Palfrey (1996), we obtain

$$\begin{aligned} W_{i\theta}(\beta, \gamma) &= E_m[U(C_i(m; \alpha, \beta, \gamma); t|t)]_{\theta=t^2/(1+V)} \\ &= \frac{V}{1+V} [1 - \beta + \beta/n + \beta\gamma\hat{\alpha}_i]^2 \\ &\quad + V \sum_{j \neq i} [\beta/n + \beta\gamma\hat{\alpha}_j]^2 \\ &\quad + \frac{\theta}{1+V} \left[1 + V \left(\frac{n-1}{n} \beta - \beta\gamma\hat{\alpha}_i \right) \right]^2, \end{aligned} \quad (1)$$

where

$$\hat{\alpha}_i = \alpha_i - \frac{1}{n}.$$

The analysis is simplified by the substitution $\mu = \beta\gamma$, which can be interpreted as the total weight of the House in the policies followed in the districts. This yields

$$\begin{aligned} W_{i\theta}(\beta, \mu) &= \frac{V}{1+V} \left[(1 - \beta) + \frac{\beta}{n} + \mu\hat{\alpha}_i \right]^2 \\ &\quad + V \sum_{j \neq i} \left[\frac{\beta}{n} + \mu\hat{\alpha}_j \right]^2 + \frac{\theta}{1+V} \left[1 + V \left(\frac{n-1}{n} \beta - \mu\hat{\alpha}_i \right) \right]^2, \end{aligned} \quad (2)$$

with the constraints $0 \leq \beta \leq 1$ and $0 \leq \mu \leq \beta$.

Voter (i, θ) prefers the confederation that minimizes $W_{i\theta}$, and level surfaces of $W_{i\theta}$ are the indiffer-

ence curves of type (i, θ) over different (β, μ) confederations.

Convexity of $W_{i\theta}$

Theorem 1 states that $W_{i\theta}$ is convex and therefore implies that the first-order conditions are sufficient to determine voter ideal points.

THEOREM 1. *For all values of the parameters, $W_{i\theta}$ is convex. It is strictly convex if and only if*

$$A_i \equiv -\hat{\alpha}_i^2 + (n-1) \sum_{j \neq i} \hat{\alpha}_j^2 > 0.$$

Proof. See the Appendix.

If we fix $\hat{\alpha}_i$, then the minimum of A_i is obtained when all the $\hat{\alpha}_j$ s are equal to one another and therefore equal to $-\hat{\alpha}_i/(n-1)$. In this case, A_i is equal to 0. This implies $A_i \geq 0$ and, furthermore, that the following corollary holds.

COROLLARY 1. *The function $W_{i\theta}$ is linear if and only if $\hat{\alpha}_j = \hat{\alpha}_{j'}$ for all j and j' not equal to i . Otherwise $W_{i\theta}$ is strictly convex.*

For most of the remainder of the article, we assume $A_i > 0$ for all districts. The case $A_i = 0$ is treated separately later.

Indifference Maps in (β, μ) Space

Except in unusual cases, the indifference curves are ellipses centered at different points of the $\mu = 0$ axis. That is, all voters' ideal points correspond to unit rather than population representation, regardless of the size of a voter's state.

THEOREM 2. *The indifference curves of the voter of type θ in district i are ellipses, centered at $(\mu^0(\theta), \beta^0(\theta))$, where*

$$\mu^0(\theta) = 0,$$

and

$$\beta^0(\theta) = \frac{1-\theta}{1+\theta V + \frac{(1-\theta)V}{n}}. \quad (3)$$

If $\hat{\alpha}_i \neq 0$, the major axes of these ellipses have slope r , defined by

$$|r| - \frac{1}{|r|} = - \frac{n \left((1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1) \hat{\alpha}_i^2 \right)}{\hat{\alpha}_i [n + V + \theta V(n-1)]} + \frac{(n-1)}{n \hat{\alpha}_i} \quad (4)$$

and

$$\text{sgn } r = \text{sgn } \hat{\alpha}_i. \quad (5)$$

For any β , the slope of the indifference curves at the point $(\beta, 0)$ is

$$-\frac{\partial W_{i\theta}/\partial \beta}{\partial W_{i\theta}/\partial \mu} = \frac{n-1}{n \hat{\alpha}_i}. \quad (6)$$

If $\hat{\alpha}_i = 0$, the ellipses have their main axis parallel to the $\beta = 0$ axis.

Proof. See the Appendix.

To summarize, if $A_i > 0$, then the slope of the main axis of the indifference curves of all voters in district i has the same sign. This slope is positive for all voters from large districts ($\hat{\alpha}_i > 0$) and negative for all voters from small districts ($\hat{\alpha}_i < 0$). Furthermore, the slope of the indifference curves of any voter is constant along the β axis. This slope only depends on the size of her district and is of the same sign as the slope of the main axis of the indifference curves.

Voters' Ideal Confederations

Since $W_{i\theta}$ is strictly convex, we can characterize the ideal points by first-order conditions. It is clear geometrically, and easily checked, that the maximum of $W_{i\theta}$ is reached at the common center of the indifference curves of the voter, $(\mu^0(\theta), \beta^0(\theta))$. It is then easy to identify the ideal feasible confederation.

THEOREM 3. *If $A_i > 0$ and $\theta < 1$, the voter's ideal feasible confederation is $(\mu^*(\theta), \beta^*(\theta)) = (0, \beta^0(\theta))$.*

If $A_i > 0$ and $\theta > 1$, then the voter's ideal feasible confederation satisfies

$$\mu^*(\theta) = \beta^*(\theta) = 0.$$

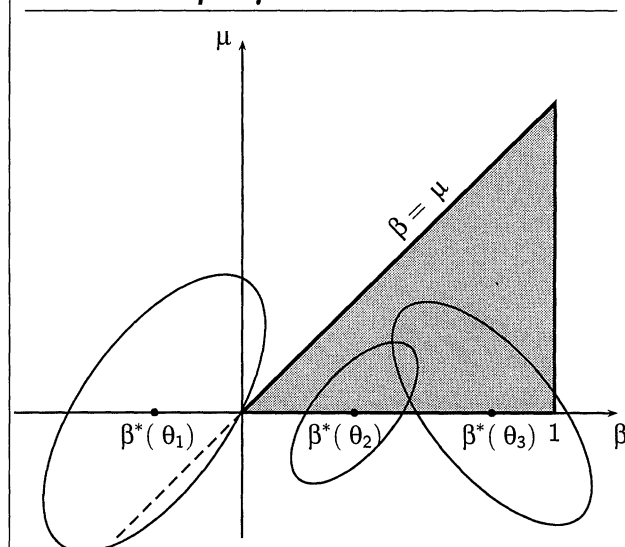
The function β^ is decreasing in θ .*

Proof. See the Appendix.

Figure 1 illustrates the feasible set and some representative indifference curves for voters from three hypothetical districts. The horizontal axis is the centralization dimension (β), and the vertical axis is the representation dimension (μ). The shaded area marks the set of feasible confederations. Voter 1 is an extreme voter from a large district. To illustrate the fact that every extreme voter's ideal confederation is $(0, 0)$, the figure includes the indifference curve of voter 1 that passes through the origin. Voters 2 and 3 are moderate voters from large and small districts, respectively.

Discussion and Comparative Statics

The first part of theorem 3 states that moderate voters ($\theta < 1$), regardless of the size of their district, are unanimously opposed to population representation. They prefer the national policy to be decided by representative institutions with the same number of delegates from each district, and they prefer power to be transferred to districts via decentralization rather than population representation. For any level of centralization, β , the variance of the centralized component is minimized by setting $\mu = 0$. Because the only

FIGURE 1. Typical Indifference Curves of Voters in the $\beta \times \mu$ Plane

Note: This figure illustrates the indifference curves of three voters. From theorem 2 and the discussion in the text, we see that $\theta_3 < \theta_2 < \theta_1$. Voter 3 is the most moderate of the three and is from a small district (the slope of the major axis of her indifference curve is negative). Voters 2 and 1 are, respectively, a less moderate voter and an extreme voter, both from large districts. Voter 1 illustrates another point. He would like a negative β , but this is infeasible. The most preferred feasible confederation for such extreme voters is always $(0, 0)$, since the indifference curve passing through $(0, 0)$ cuts the diagonal from below.

advantage of centralization is risk reduction, it follows that any voter's ideal confederation must have $\mu = 0$.

The second half of the theorem is even simpler to understand. For extremist voters ($\theta > 1$), the expression obtained for β^* from the first-order conditions is negative, so the feasibility constraints bind. They want no centralization at all, which corresponds to $\beta = 0$. At $\beta = 0$, the only feasible value of μ is 0.

Thus, we have established that the optimal confederation for all voters from all districts calls for no population representation at the national level. We should keep in mind, however, that this is not the same as saying that all voters *always* prefer unit representation (this point is discussed in more detail in the next subsection).

The comparative statics of how the ideal points change with respect to the exogenous parameters of the model are straightforward. For voters with $\theta < 1$, it is easy to see that $\beta^*(\theta)$ is increasing in n . The intuition behind this is simple. The moderation, or risk-reduction, benefits of centralization are greater as n is greater, independent of district size, since having more districts reduces the probability that the centralized policy will be dominated by an extremist majority of one wing or the other. For similar reasons, $\beta^*(\theta)$ is decreasing in V . The higher is V , the more dispersed is the distribution of voter ideal points within a district, so there is a higher probability that the centralized policy will be on one extreme or the other.

Conditional Ideal Points

The final question we address is a voter's optimal value of μ as a function of β : For a given degree of centralization, how much power does a voter want the House to have? The answer will be useful later in the article, when we study the properties of majority rule equilibrium.

For voters with $\hat{\alpha}_i \neq 0$, the optimal value of μ generally depends on the choice of β . This relationship, $\mu^*(\beta)$, depends on the orientation of the (ellipsoidal) indifference curves of the voters in (β, μ) space. To find the solution, consider again the first-order condition from differentiation of W with respect to μ , holding β fixed:

$$(1 - \theta)\hat{\alpha}_i - \beta \frac{\hat{\alpha}_i}{n} [n + V + \theta V(n - 1)] + \mu \left[(1 + V) \sum_{j=1}^n \hat{\alpha}_j^2 - (1 - \theta)V\hat{\alpha}_i^2 \right] = 0.$$

We solve for μ and summarize the results in the following theorem.

THEOREM 4. Fix β . The optimal ideal confederation for a voter θ from district i is

$$\mu^*(\beta, \theta) = -\hat{\alpha}_i \left[\frac{(1 - \theta) - \frac{\beta}{n} [n + V + \theta V(n - 1)]}{(1 + V) \sum_{j=1}^n \hat{\alpha}_j^2 - (1 - \theta)V\hat{\alpha}_i^2} \right] \quad (7)$$

if

$$0 \leq -\hat{\alpha}_i \left[\frac{(1 - \theta) - \frac{\beta}{n} [n + V + \theta V(n - 1)]}{(1 + V) \sum_{j=1}^n \hat{\alpha}_j^2 - (1 - \theta)V\hat{\alpha}_i^2} \right] \leq \beta.$$

Otherwise, $\mu^*(\beta, \theta)$ equals either 0 or β , depending on whether the right-hand side of equation 7 is less than 0 or greater than β , respectively.

Proof. See the Appendix.

Whether $\mu^*(\beta, \theta)$ is greater than 0 depends on whether $\hat{\alpha}_i$ is greater than 0 (large districts) or less than 0 (small districts) and whether β is greater than or less than $\beta^*(\theta)$.

Consider, first, the case of large districts. Since $\hat{\alpha}_i > 0$, the sign of $\mu^*(\beta, \theta)$ is the same as the sign of

$$(1 - \theta) - \beta \left[1 + \theta V + \frac{(1 - \theta)V}{n} \right] - \frac{(1 + V) \sum_{j \neq i} \hat{\alpha}_j^2 + (1 + \theta V)\hat{\alpha}_i^2}{(1 + V) \sum_{j \neq i} \hat{\alpha}_j^2 + (1 + \theta V)\hat{\alpha}_i^2}.$$

Since the denominator is positive, $\mu^*(\beta, \theta) > 0$ for $\theta > 1$. This establishes that all extremists in larger states

prefer at least some degree of population representation for every value of $\beta > 0$.¹¹

More generally, in large districts, $\mu^*(\beta, \theta) > 0$ if and only if $\beta > \beta^*(\theta)$. That is, voters from large districts favor some degree of population representation if and only if there is too much centralization relative to their preferred level. The intuition is that by putting some weight on population representation, the large district can effectively recoup some of the sovereignty lost because of a high β . Voters from large districts who are sufficiently extreme want to do this. This can also be interpreted in the following way. Let $\theta^*(\beta)$ be the solution of $\beta^*(\theta) = \beta$. That is, $\theta^*(\beta)$ is the voter for whom β is the ideal level of centralization. The critical value $\theta^*(\beta)$ is decreasing in β , so voters in large districts prefer some degree of population representation if and only if $\theta < \theta^*(\beta)$. In other words, voters in large districts who are more extreme than $\theta^*(\beta)$ would like some population representation. The critical value $\theta^*(\beta)$ is decreasing in β .

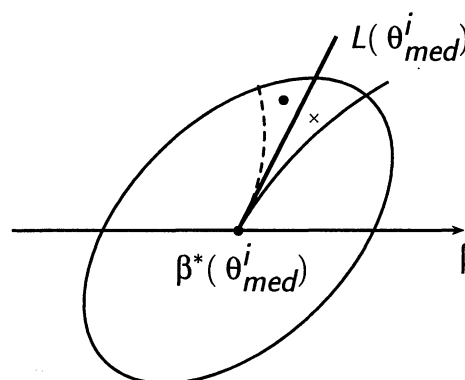
Not surprisingly, preferences over β in small districts go in exactly the opposite direction. Voters from small districts favor some degree of population representation if and only if $\beta < \beta^*(\theta)$. The intuition here is exactly the reverse of the previous intuition. In this case, relatively moderate voters (i.e., voters for whom $\beta^*(\theta) > \beta$) want to cede even more sovereignty. This is done by increasing the weight on population representation, since doing so reduces the influence of a small district on the centralized component of the policy. For any fixed value of β there is a critical voter $\theta^*(\beta)$, such that voters from small districts prefer some degree of a population representation if and only if $\theta > \theta^*(\beta)$.

One can conduct a similar analysis of conditional ideal points along the centralization dimension, fixing μ . When $\mu = 0$, a voter's conditional ideal degree of centralization does not depend on state size, as established in the previous sections. However, if $\mu > 0$, so there is at least some mixture of population and unit representation, then this is no longer the case; voters in large states will prefer more centralization than otherwise identical voters in small states. This is easy to see from inspection of Figure 1. Similarly, the greater is μ , the more centralization voters in a large state desire, while the opposite is true for voters in small states.

At this point, we can speculate that the qualitative features of the results we have obtained are robust to a variety of other specifications beyond our normal-quadratic model. For instance, as explained in Cr  mer and Palfrey (1996), it should be the case very generally that moderate voters prefer centralization. Furthermore, the intuition behind the results on the relationship between district size and preferences over representation and centralization, as discussed in the three previous paragraphs, seems quite general.

¹¹ This is also straightforward geometrically. Consider voter 1 in Figure 1, and fix β to some strictly positive value, β' . The tangency between the vertical line $\beta = \beta'$ and an indifference curve will occur for some strictly positive value of μ .

FIGURE 2. Informal Proof of Theorem 5



Note: The point marked $\beta^*(\theta_{med}^i)$ is the median ideal confederation in district i . It is also the majority voting equilibrium in district i . To show this, we have graphed the indifference curves through $\beta^*(\theta_{med}^i)$ of a typical voter with a small θ (the solid curve) and a typical voter with a large θ (the dashed curve). Any move to a confederation such as the point marked \times will be voted down by all the voters with $\theta > \theta_{med}^i$ and some with $\theta < \theta_{med}^i$. A move to a point such as the one marked with a dot will be voted down by all the voters with $\theta < \theta_{med}^i$ and some with $\theta > \theta_{med}^i$.

MAJORITY RULE EQUILIBRIUM

Majority Rule Equilibrium within a District

A *majority rule equilibrium* in district i is a centralization-representation pair (β, μ) with the property that there does not exist another pair (β', μ') such that a majority of members in district i strictly prefer (β', μ') to (β, μ) . One can show that every district has a majority rule equilibrium, $(\beta_i^*, 0)$, where $\beta_i^* = \beta^*(\theta_{med}^i)$, and θ_{med}^i is the median value of θ in district i . Moreover, if $A_i > 0$, then this is the unique majority rule equilibrium in that district. This is stated formally in theorem 5 and proved in the Appendix, with an informal proof and explanation in Figure 2.

THEOREM 5. *If $A_i > 0$, then $(\beta^*(\theta_{med}^i), 0)$ is the unique majority rule equilibrium confederation within district i .*

The within-district majority rule equilibrium, $\beta^*(\theta_{med}^i)$, is independent of the relative size of the district, α_i , but varies systematically with the underlying parameters of the model, n and V . Specifically, $\beta^*(\theta_{med}^i)$ is increasing in n and decreasing in V , and it is larger the more moderate is the district's median voter ideal point.

Equilibrium Confederations

The analysis in the previous section does not address the aggregation of the induced preferences across the various districts to reach an interdistrict decision about the degree of centralization and the extent of population representation. This aggregation is central to

understanding how and why different systems of representation and centralization may arise, so we address it here in the context of an equilibrium model. We identify the range of possible equilibrium confederations and consider two different concepts of majority rule equilibrium.

First, we consider majority rule equilibrium representation, conditional on some fixed level of centralization. We show how this equilibrium analysis implies that mixtures of population and unit representation (such as the House-Senate bicameral system in the United States) are likely to arise, particularly when preferences and district size are correlated in certain ways. Second, we look at unconditional equilibrium, which involves majority rule equilibrium in the two-dimensional space of representation \times centralization. As is typically the case in more than one dimension, a majority rule equilibrium does not always exist, and nonexistence is particularly a problem in precisely those preference configurations where conditional equilibrium involves a mixture of population and unit representation. Thus, it is not surprising that we often observe a combination of these two kinds of representation in real-world institutions.

Conditional Equilibrium

Here we separate out the issues of voting over the level of centralization, β , and voting over the representation structure, μ . As we see in the next subsection, it is often necessary to separate these two dimensions in order to guarantee the existence of a majority rule equilibrium. Thus, for any $\beta \in [0, 1]$, we define a β -conditional equilibrium confederation to be a representation structure,¹² μ , with the property that there does not exist another representation structure, μ' , such that a majority of voters in the confederation prefer (β, μ') to (β, μ) . So a conditional equilibrium confederation is simply one against which any alternative confederationwide referendum regarding the representation dimension—with every voter in every district voting—would fail to win a majority of votes, fixing the level of centralization.¹³ Our first result is that for every $\beta \in [0, 1]$ there exists a unique β -conditional equilibrium confederation.

THEOREM 6. *If $N > 2$, then for every $\tilde{\beta} \in [0, 1]$ there exists a unique value of $\mu \in [0, 1]$, denoted $\mu^*(\tilde{\beta})$, such that $(\tilde{\beta}, \mu^*(\tilde{\beta}))$ is a $\tilde{\beta}$ -conditional equilibrium confederation.*

Proof. As established earlier, voters have unique conditional ideal points for any fixed centralization param-

eter, $\tilde{\beta}$. Theorem 4 implies that voters have single-peaked preferences over $\{(\beta, \mu) | \beta = \tilde{\beta} \text{ and } \mu \in [0, 1]\}$. Therefore, by the median voter theorem, there is a $\tilde{\beta}$ -conditional equilibrium confederation at $\mu^*(\tilde{\beta}) = \mu_m^*(\tilde{\beta}, \theta_m)$ where m is the median conditional ideal point given $\tilde{\beta}$. Q.E.D.

A natural question, then, is whether $\mu^*(\tilde{\beta}) > 0$ is either possible or probable. The answer is yes, but this depends on both the configuration of voter ideal points and the location of the median voter in states of different sizes.

For simplicity, consider the following simple but robust example. There are three districts whose median values of θ are given by θ_1 , θ_2 , and θ_3 , all strictly positive. We suppose $\alpha_3 < \alpha_2 < 1/3 < \alpha_1$, and $\beta^*(\theta_1) < \beta(\bar{\theta}) < \beta^*(\theta_2) < \beta^*(\theta_3)$, where $\bar{\theta} = \alpha_1\theta_1 + \alpha_2\theta_2 + \alpha_3\theta_3$. We will show that for any level of centralization $\tilde{\beta}$ in the interval $(\beta^*(\theta_1), \beta^*(\theta_2))$ (which includes $\beta(\bar{\theta})$), the $\tilde{\beta}$ -conditional equilibrium confederation, $\mu^*(\tilde{\beta})$, is strictly positive.

Indeed, by theorem 3, β^* is decreasing. Hence, for all $\theta > \theta_1$, we have $\beta^*(\theta) < \beta^*(\theta_1)$. By continuity of preferences, for strictly more than half of the voters in district 1, $\beta^*(\theta) < \tilde{\beta}$, and by the discussion that follows theorem 4, for all these voters $\mu^*(\tilde{\beta}, \theta) > 0$, because district 1 is large.

Similarly, for $i = 2, 3$, for all $\theta < \theta_i$, we have $\beta^*(\theta) > \beta^*(\theta_i)$. By continuity of preferences, for strictly more than half the voters in district i , $\beta^*(\theta_i) > \tilde{\beta}$, and by the discussion that follows theorem 4, for all these voters $\mu^*(\tilde{\beta}, \theta) > 0$, because districts 2 and 3 are small.

Hence, in all districts there is a majority of voters for whom $\mu^*(\tilde{\beta}) > 0$, and we have proved the result. Furthermore, by continuity, if either $\beta^*(\theta_2)$ is sufficiently close to $\beta^*(\theta_3)$ or district 2 is sufficiently smaller than district 3, then for any $\tilde{\beta}$ between $\beta^*(\theta_1)$ and $\beta^*(\theta_3)$, the $\tilde{\beta}$ -conditional equilibrium confederation is strictly positive. Since reasonable methods of choosing a constitution can be expected to yield $\tilde{\beta}$ in the interval $[\beta^*(\theta_1), \beta^*(\theta_3)]$ (that is, between the minimum and maximum median district ideal degrees of centralization), the $\tilde{\beta}$ -conditional equilibrium confederation would always involve a mixture of the two types of representation.

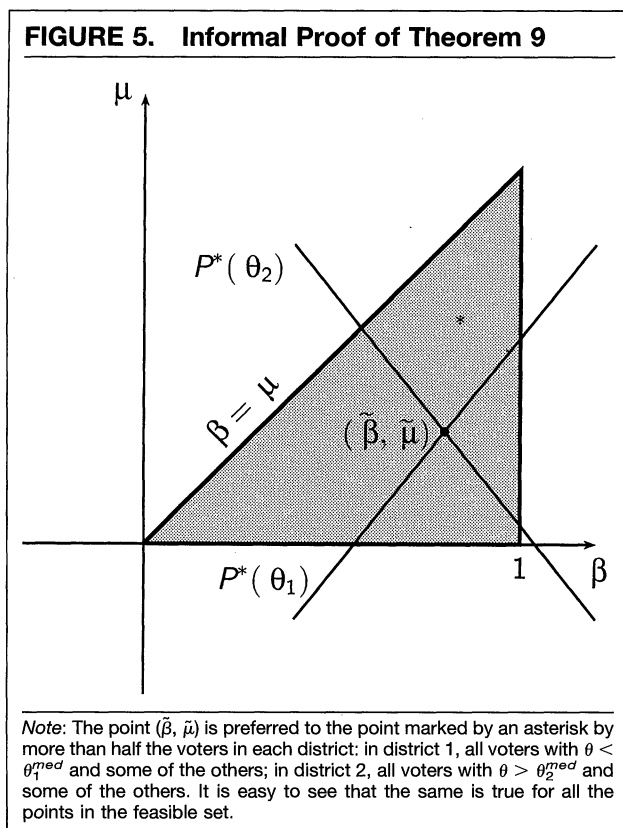
Unconditional Equilibrium

An unconditional equilibrium confederation is a centralization-representation pair (β, μ) with the property that there does not exist another pair (β', μ') such that a majority of voters in the confederation prefer (β', μ') to (β, μ) . So an unconditional equilibrium confederation is simply one against which any alternative confederationwide referendum—with every voter in every district voting—would fail to win a majority of votes. It is unconditional because we permit the referendum to change both the degree of centralization and the mixture of representation simultaneously.

We can characterize the equilibrium as follows.

¹² A μ -conditional equilibrium confederation for any $\mu \in [0, 1]$ can be defined in a similar fashion, but we limit our formal analysis here to the first kind of conditional equilibrium.

¹³ There are other possible ways to aggregate votes at the confederation level, including, for example, a convention to which each district sends a delegate, who is instructed to vote according to the majority preference within his district. Theorems 6–10 hold for this kind of vote aggregation, too.



Proof. See the Appendix.

When $n = 2$, $A_1 = A_2 = 0$, and we can show the following theorem.

THEOREM 9. *If $n = 2$, then there exists a unique majority rule equilibrium.*

Proof. See the Appendix. An informal proof is given in Figure 5.

This equilibrium may lie anywhere between the β axis and the $\beta = \mu$ diagonal.

In the special case in which $\alpha_1 = \alpha_2 = 1/2$, all confederations with the same β produce equivalent policies in every district, so uniqueness is defined relative to such equivalence classes of confederations.

More generally, if all districts are of the same size, that is, if $\alpha_i = 1/n$ for all i , then the centralized policy is neutral with respect to μ , since the population and unit representation models are equivalent. Majority rule equilibria exist and are of the form $(\beta^*(\bar{\theta}), \mu)$, where $\bar{\theta}$ is the median θ and μ is any value between 0 and 1.

CONCLUSION

This article investigated questions of sovereignty and representation in confederations. We model sovereignty formally by allowing policy outcomes in one district to be affected by the distribution of voters in other districts and by the relative size of districts. We model the representation dimension as a choice of the

degree to which each district is represented proportionally to population as opposed to unit based. The balance of power between the two houses of the U.S. Congress is a classic example of how such trade-offs may be achieved in practice. The questions are posed at both the individual and aggregate level.

At the individual level, the questions reduce to the following.

- (I1) To what extent do voters prefer to sacrifice sovereignty of their own district in order to ensure more moderate policy outcomes?
- (I2) How do voters in large states differ from voters in small states in terms of induced preferences over population versus unit representation?

The answer to I1 generalizes the earlier findings in Cr mer and Palfrey (1996). Induced preferences over sovereignty cut across traditional Left-Right cleavages. Instead, extremists from both sides of the political spectrum are in agreement and in opposition to relatively moderate voters, who prefer more centralization. Sufficiently extreme voters are "separatists," who ideally prefer no confederation at all. We also show that preference for more centralization (i.e., willingness to cede sovereignty) increases with n , decreases with V , and is independent of the size distribution of districts.

The answer to I2 is more complicated. For all voters, the ideal confederation would consist entirely of unit representation, but preferences over representation cannot be disentangled from issues of centralization. That is, voters' induced preferences over population versus unit representation depend on the degree of centralization. When there is too much centralization, voters in large states would like some degree of population representation, which restores some local sovereignty. The opposite is true for voters in small states, who would like some degree of population representation only if there is too little centralization.

At the aggregate level, there are two central questions.

- (A1) What are the properties of majority rule equilibrium in a district?
- (A2) What are the properties of majority rule equilibrium across the whole federation?

The answer to A1 is that a majority rule equilibrium exists within each district and coincides with the ideal confederation of the median value of θ in that district. The answer to A2 depends on whether one considers conditional or unconditional majority rule equilibrium. In a conditional equilibrium, for any fixed level of centralization, there is always a unique (conditional) majority rule equilibrium in the representation dimension. This conditional equilibrium could correspond to cases in which the centralization dimension has already been settled, in which there is relatively little flexibility on the centralization dimension, or in which the agenda has been organized in a way that separates the two dimensions. In the unconditional equilibrium, the two dimensions are voted on simultaneously. As is

typical in such cases, a majority rule equilibrium may fail to exist due to majority voting cycles.

In the conditional equilibrium, the representation dimension typically involves a mixture of population and unit representation, consistent with what one observes in a great many representative legislatures. The properties of conditional ideal points of voters suggest the kinds of coalitions that may form in a vote along the representation dimension. *Ceteris paribus*, moderate voters in small states and extreme voters in large states will push for greater weight on population representation, while moderates in large states and extremists in small states will want to shift more weight to unit representation.

The properties of unconditional equilibrium depend on whether there are two districts or more than two. With two districts, a unique equilibrium confederation always exists. Depending on the relative extremism of the median voters in the two states, the equilibrium degree of population representation can range from entirely unit representation to entirely population representation. With three or more districts, equilibrium may fail to exist. When it does exist, representation is entirely unit based.

The possibility of majority rule cycling in the unconditional equilibrium underscores the importance of agenda manipulation and procedural control at constitutional conventions. The empirical relevance of such manipulation has been noted in historical studies of constitutional conventions. For example, Riker (1984, 1986) recounts numerous episodes of agenda manipulation at the U.S. Constitutional Convention of 1787, consistent with the sort of preference aggregation problem identified in this article. This suggests that an exploration of agenda effects in our model of confederations would be an interesting direction for further research.

APPENDIX: FORMAL PROOFS OF THEOREMS

Proof of Theorem 1

The function W_{i0} is convex if and only if the determinant

$$D_{i0} = \begin{vmatrix} \partial^2 W_{i0} / \partial \beta^2 & \partial^2 W_{i0} / \partial \beta \partial \mu \\ \partial^2 W_{i0} / \partial \beta \partial \mu & \partial^2 W_{i0} / \partial \mu^2 \end{vmatrix}$$

is nonnegative. W is strictly convex if and only if $D_{i0} > 0$. Expanding W_{i0} gives

$$\begin{aligned} W_{i0}(\beta, \mu) = & \frac{V}{1+V} \left[1 + \beta^2 \left(1 - \frac{1}{n} \right)^2 + \mu^2 \hat{\alpha}_i^2 \right. \\ & - 2\beta \left(1 - \frac{1}{n} \right) + 2\mu \hat{\alpha}_i - 2\beta \mu \left(1 - \frac{1}{n} \right) \hat{\alpha}_i \\ & + V \sum_{j \neq i} \left[\frac{\beta^2}{n^2} + \mu^2 \hat{\alpha}_j^2 + 2 \frac{\beta}{n} \mu \hat{\alpha}_j \right] \\ & \left. + \frac{\theta}{1+V} \left[1 + V^2 \left(1 - \frac{1}{n} \right)^2 \beta^2 + V^2 \mu^2 \hat{\alpha}_i^2 \right] \right] \end{aligned}$$

$$\begin{aligned} & + 2\beta V \left(1 - \frac{1}{n} \right) - 2\mu V \hat{\alpha}_i - 2V^2 \beta \mu \left(1 - \frac{1}{n} \right) \hat{\alpha}_i \\ & = \frac{\beta^2 V(n-1)}{n^2 (1+V)} [n + V + V\theta(n-1)] \\ & + \mu^2 \frac{V}{1+V} \left[(1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1) \hat{\alpha}_i^2 \right] \\ & - 2\beta \mu \frac{V}{1+V} \frac{\hat{\alpha}_i}{n} [n + V + \theta V(n-1)] \\ & + 2\mu \frac{V}{1+V} \hat{\alpha}_i [1 - \theta] \\ & - 2\beta \frac{V}{1+V} \frac{(n-1)}{n} [1 - \theta] + \frac{V + \theta}{1+V}. \end{aligned}$$

Taking partial derivatives gives

$$\begin{aligned} & \frac{(1+V)^2}{V^2} D_{i0} \\ & = \begin{vmatrix} \frac{n-1}{n^2} (n + V + V\theta(n-1)) & -\frac{\hat{\alpha}_i}{n} [n + V + \theta V(n-1)] \\ -\frac{\hat{\alpha}_i}{n} [n + V + \theta V(n-1)] & (1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1) \hat{\alpha}_i^2 \end{vmatrix} \\ & = \frac{n + V + V\theta(n-1)}{n^2} \begin{vmatrix} n-1 & -\hat{\alpha}_i [n + V + \theta V(n-1)] \\ -\hat{\alpha}_i & (1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1) \hat{\alpha}_i^2 \end{vmatrix} \end{aligned}$$

Hence, D_{i0} has the same sign as

$$\begin{aligned} & (n-1) \left((1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1) \hat{\alpha}_i^2 \right) \\ & - \hat{\alpha}_i^2 (n + V + \theta V(n-1)) = (1+V) A_i. \end{aligned}$$

Proof of Theorem 2

To show that our quadratic function defines an ellipse, it is sufficient to show that the function is strictly convex, which is guaranteed by theorem 1. The center is found by setting the derivatives equal to 0, which yields

$$(1 - \theta) - \left[\frac{\beta}{n} - \frac{\mu \hat{\alpha}_i}{n-1} \right] [n + V + \theta V(n-1)] = 0, \quad (9)$$

through derivation with respect to β , and

$$\begin{aligned} & (1 - \theta) \hat{\alpha}_i - \beta \frac{\hat{\alpha}_i}{n} [n + V + \theta V(n-1)] \\ & + \mu \left[(1+V) \sum_{j=1}^n \hat{\alpha}_j^2 - (1 - \theta) V \hat{\alpha}_i^2 \right] = 0, \quad (10) \end{aligned}$$

through derivation with respect to μ . Multiplying equation 9 by $-\hat{\alpha}_i$ and adding it to equation 10, we obtain

$$\mu \left[(1+V) \sum_{j=1}^n \hat{\alpha}_j^2 - V \hat{\alpha}_i^2 - \frac{\hat{\alpha}_i^2}{n-1} (n + V) \right] = 0.$$

Since $A_i > 0$, the term in brackets is not equal to 0, so $\mu^0(\theta) = 0$, and we find $\beta^0(\theta)$ by substituting $\mu = 0$ into equation 9.

To see that the absolute value of r satisfies equation 4, we apply the standard formula for the slope of an ellipse, which is

$$r - \frac{1}{r} = \frac{A - B}{C},$$

where A is the coefficient on β^2 , B is the coefficient on μ^2 , and C is the coefficient on $\beta\mu$ in the quadratic expression for $W_{i\theta}(\beta, \mu)$. Note that $r - 1/r$ is increasing on $(-\infty, 0)$ and $(0, +\infty)$ and takes values on both these intervals on $(-\infty, +\infty)$. Therefore, the formula only determines r up to its sign. The formula gives

$$r - \frac{1}{r}$$

$$\begin{aligned} & (1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1)\hat{\alpha}_i^2 - \frac{n-1}{n^2} [n+V+V\theta(n-1)V\theta] \\ &= \frac{-\frac{\hat{\alpha}_i}{n} [n+V+V\theta(n-1)]}{n \left((1+V) \sum_{j=1}^n \hat{\alpha}_j^2 + V(\theta-1)\hat{\alpha}_i^2 \right)} + \frac{n-1}{n\hat{\alpha}_i} \\ &= \frac{1}{\hat{\alpha}_i} \left(\frac{n-1}{n} - \frac{n((1+V) \sum_{j \neq i} \hat{\alpha}_j^2 + (1+V\theta)\hat{\alpha}_i^2)}{n+V+V\theta(n-1)} \right). \end{aligned}$$

The sign of r is the opposite of the sign of A/C and hence the sign of $\hat{\alpha}_i$.

To prove equation 6, observe that

$$\begin{aligned} & -\frac{\partial W_{i\theta}/\partial \beta}{\partial W_{i\theta}/\partial \mu} \Big|_{\mu=0} \\ &= -\frac{\frac{(n-1)V}{n(1+V)} \left[-1 + \theta + \frac{\beta}{n} (n+V+V\theta(n-1)) \right]}{\frac{V}{1+V} \left[\hat{\alpha}_i(1-\theta) - \frac{\beta}{n} \hat{\alpha}_i(n+V+V\theta(n-1)) \right]} \\ &= -\frac{n-1}{n} \times \frac{1}{-\hat{\alpha}_i}. \end{aligned}$$

Proof of Theorem 3

If $\theta < 1$, then $\beta^0(\theta) \in [0, 1]$, and the feasibility constraints ($\beta \in [0, 1]$ and $\mu \in [0, \beta]$) do not bind. The solution given by the first-order conditions characterizes the minimum value of W over the set of feasible confederations. Hence,

$$\mu^* = 0,$$

and

$$\beta^* = \frac{n(1-\theta)}{n+V+(n-1)\theta V}.$$

If $\theta > 1$, then the unconstrained value of β^* is negative, so at least one constraint ($\beta \in [0, 1]$ or $\mu \in [0, \beta]$) must bind. From equation 6, the slope of the indifference curve at $(0, 0)$ is either negative or strictly greater than 1. Therefore, $(0, 0)$ is a local maximum in the feasible set and, by convexity, a global maximum.

To see that β^* is decreasing in θ , simply rewrite $\beta^0(\theta)$ as

$$\frac{1-\theta}{1+V/n+\theta V(1-1/n)}$$

and notice that the numerator is decreasing in θ while the denominator is increasing.

Proof of Theorem 4

Solving equation 10 for μ , we get

$$\mu = -\frac{(1-\theta)\hat{\alpha}_i - \beta \frac{\hat{\alpha}_i}{n} [n+V+V\theta(n-1)]}{(1+V) \sum_{j=1}^n \hat{\alpha}_j^2 - (1-\theta)V\hat{\alpha}_i^2}$$

If this μ satisfies the feasibility constraints, then it is equal to μ^* . Otherwise, by convexity of W , the solution lies on the boundary, as described in the theorem.

Proof of Theorem 5

First notice that if $\beta^*(\theta_{med}^i) \leq 0$, then the result is trivial, as $(0, 0)$ is the most preferred point in the feasible set for more than half the voters (those for whom $\beta_i^*(\theta) \leq 0$).

Thus, suppose $\beta^*(\theta_{med}^i) > 0$. Consider in Figure 2 the two half-planes defined by the line $L(\theta_{med}^i)$, following the major axis of the indifference map of voter θ_{med}^i and passing through $(\beta^*(\theta_{med}^i), 0)$. Consider any point in the half-plane consisting of all points to the left of this line, such as the dot in the figure. Since the major axes of the voters are parallel to each other along the β -axis, it follows immediately that all voters with θ such that $\beta_i^*(\theta) > \beta^*(\theta_{med}^i)$ prefer $\beta^*(\theta_{med}^i)$ to y . This is, by definition of $\beta^*(\theta_{med}^i)$, exactly half the voters, so that y does not defeat $\beta^*(\theta_{med}^i)$ by a majority. A similar argument applies to any point in the right half-plane, such as the point marked with a cross in the figure. In this case, all voters with θ such that $\beta_i^*(\theta) < \beta^*(\theta_{med}^i)$ prefer $\beta^*(\theta_{med}^i)$ to y , so y does not defeat $\beta^*(\theta_{med}^i)$ by a majority.

To show uniqueness of the majority rule equilibrium, notice that for any point (β', μ') not equal to $(\beta^*(\theta_{med}^i), 0)$, continuity implies that all voters with ideal points in a neighborhood of $(\beta(\theta_{med}^i), 0)$ prefer it to (β', μ') . By the argument above, it is also true that either all voters to the left of θ_{med}^i or all voters to the right of θ_{med}^i prefer $(\beta^*(\theta_{med}^i), 0)$ to (β', μ') . Thus, a strict majority prefers $(\beta^*(\theta_{med}^i), 0)$ to (β', μ') , which cannot be a majority rule equilibrium.

Proof of Theorem 7

By corollary 1, under the conditions of the theorem, $A_i > 0$ for all i , and therefore all indifference curves are ellipses.

We first show $\tilde{\mu} = 0$. Suppose not, and consider an equilibrium $e = (\tilde{\beta}, \tilde{\mu})$, with $\tilde{\mu} \in (0, \tilde{\beta}]$ and $\tilde{\beta} > 0$. Denote by $\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i)$ the optimal β for a type θ voter from a district with relative size $\hat{\alpha}_i$, given $\tilde{\mu}$. Formally, this is obtained by rearranging the first-order condition, equation 4, to get:

$$\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i) = \frac{n(1-\theta)}{n+V+(n-1)\theta V} + \frac{n\tilde{\mu}\hat{\alpha}_i}{n-1}.$$

Note that this may fall outside the feasible bounds given by the restriction that $\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i) \in [\tilde{\mu}, 1]$. Since we have supposed that $(\tilde{\beta}, \tilde{\mu})$ is a majority rule equilibrium, it must be the case that $\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i) > \tilde{\beta}$ for 50% of the voters and $\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i) < \tilde{\beta}$ for 50% of the voters. If not, then one could either increase or decrease $\tilde{\beta}$ slightly, keeping $\tilde{\mu}$ fixed at $\tilde{\mu}$, and a majority would vote for the new proposal. Now consider all voters for whom $\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i) < \tilde{\beta}$. A typical voter like this is shown in Figure 3, where we denote $\beta_{\tilde{\mu}}^*(\theta, \hat{\alpha}_i)$ by v^2 .

Now consider rotating the horizontal line, defined by $P =$

$\{(\beta, \mu) | \mu = \bar{\mu}\}$ clockwise by a very small angle of rotation, ϕ . (If $\bar{\mu} = \bar{\beta}$, then rotate the horizontal line counterclockwise.) This new (downward-sloping) line, denoted PQ , is shown in the figure. For each voter type $(\theta, \hat{\alpha}_i)$, define $p_{P_\phi}^*(\theta, \hat{\alpha}_i)$ as this voter's most preferred feasible confederation, restricted to PQ . The point, $p_{P_\phi}^*(\theta, \hat{\alpha}_i)$, for the typical voter described above is marked v^3_ϕ in the figure. Because the preference gradient for this voter points directly downward at v^2 , this implies that v^3_ϕ lies to the left of e . This is also true for any voter for whom $\beta_{\bar{\mu}}^*(\theta, \hat{\alpha}_i) = \bar{\beta}$, as illustrated by the point labelled v^1 in the figure, and by continuity it is also true for a positive measure of voters for whom $\beta_{\bar{\mu}}^*(\theta, \hat{\alpha}_i) < \bar{\beta}$, such that $\beta_{\bar{\mu}}^*(\theta, \hat{\alpha}_i)$ is sufficiently close to $\bar{\beta}$. Thus, there must exist a proposal lying to the left of e on P_ϕ that a majority of voters prefer to $(\bar{\beta}, \bar{\mu})$. Therefore, $(\bar{\beta}, \bar{\mu})$ cannot be a majority rule equilibrium if $\bar{\mu} > 0$.

Finally, suppose that $(\bar{\beta}, 0)$ is a majority rule equilibrium, where $\bar{\beta} \neq \beta^*(\bar{\theta})$. Then there exists a majority of voters such that either $\beta^*(\bar{\theta}) < \bar{\beta}$ or $\beta^*(\bar{\theta}) > \bar{\beta}$. Suppose without loss of generality there is a majority of voters such that $\beta^*(\bar{\theta}) > \bar{\beta}$. Then, for small enough ε , for all $\beta' \in (\bar{\beta}, \bar{\beta} + \varepsilon)$, a majority would prefer $(\beta', 0)$ to $(\bar{\beta}, 0)$. Thus, a necessary condition for $(\bar{\beta}, \bar{\mu})$ to be an equilibrium is that $\bar{\beta} = \beta^*(\bar{\theta})$ and $\bar{\mu} = 0$.

Proof of Theorem 8

The coefficient of μ^2 in equation 8 is equal to

$$\frac{V}{1+V} \left[(1+V) \sum_{j \neq i} \hat{\alpha}_j^2 + (1+V\theta) \hat{\alpha}_i^2 \right].$$

If $A_i = 0$, this is equal to

$$\begin{aligned} \frac{V\hat{\alpha}_i^2}{1+V} \left[\frac{1+V}{n-1} + 1 + V\theta \right] \\ = \frac{V\hat{\alpha}_i^2}{(n-1)(1+V)} [V + n + V\theta(n-1)]. \end{aligned}$$

This enables us to rewrite equation 8 as follows:

$$\begin{aligned} W_{i0}(\beta, \mu) = \frac{V(n-1)}{1+V} [n + V + V\theta(n-1)] \left(\frac{\beta}{n} - \frac{\mu\hat{\alpha}_i}{(n-1)} \right)^2 \\ - 2 \frac{V}{1+V} (n-1) [1 - \theta] \left(\frac{\beta}{n} - \frac{\mu\hat{\alpha}_i}{(n-1)} \right) + \frac{V + \theta}{1+V}, \end{aligned}$$

which shows that W_{i0} is a quadratic function of $\beta/n - \mu\hat{\alpha}_i/(n-1)$. The indifference curve associated with the minimum of W_{i0} will therefore be a single straight line of equation $\beta/n - \mu\hat{\alpha}_i/(n-1) = K$ for some K , and the indifference curves associated to greater values of W_{i0} will be pairs of such straight lines.

Proof of Theorem 9

If $\alpha_1 = \alpha_2 = 1/2$, then all voters have vertical indifference lines. Consider the ideal line of voter $\bar{\theta} = 1/2 \theta_1 + 1/2 \theta_2$, where θ_1 and θ_2 are the medians of district 1 and district 2, respectively. Any point $(\beta^*(\bar{\theta}), \mu)$ such that $\mu \in [0, \beta]$ is an equilibrium.

Next suppose that $\alpha_1 > \alpha_2$. Let $P^*(\theta_1)$ and $P^*(\theta_2)$ denote the ideal lines of θ_1 and θ_2 , respectively. It is easily verified that these two lines have slopes that are equal in magnitude (greater than 1), with opposite signs; $P^*(\theta_1)$ is upward sloping, and $P^*(\theta_2)$ is downward sloping.

There are three cases to consider, depending on the position of the intersection of $P^*(\theta_1^{med})$ and $P^*(\theta_2^{med})$:

Case 1: The intersection lies at some point $(\bar{\beta}, \bar{\mu})$, with $\bar{\mu} \in (0, \bar{\beta})$.

Case 2: It lies below the horizontal axis.

Case 3: It lies above the diagonal.

Case 1 is easiest because, for any $(\beta, \mu) \neq (\bar{\beta}, \bar{\mu})$, less than 50% of the voters prefer (β, μ) to $(\bar{\beta}, \bar{\mu})$. Thus, $(\bar{\beta}, \bar{\mu})$ is a majority rule equilibrium. Obviously, no other point can be a majority rule equilibrium in this case. This is illustrated by Figure 5.

In case 2, consider the indifference line of the voter with the median preference $\bar{\theta} = \alpha_1 \theta_1^{med} + \alpha_2 \theta_2^{med}$. For voters of this type, the median ideal confederation, $\beta = \beta^*(\bar{\theta})$, lies on the horizontal axis ($\mu = 0$), regardless of whether they are in the large or small district. Since the intersection of $P^*(\theta_1)$ and $P^*(\theta_2)$ lies below the horizontal axis, a majority of district 1 voters have ideal lines to the right of the $\bar{\theta}$ ideal line, and a majority of district 2 voters have ideal lines to the left of the $\bar{\theta}$ ideal line. Therefore, $(\beta^*(\bar{\theta}), 0)$ is a majority rule equilibrium. It is easy to verify that there is no other (β, μ) that can be an equilibrium.

Case 3 is slightly more complicated. Consider the preferences of voters restricted to (β, μ) on the diagonal. The ideal policy on this diagonal for a θ_1^{med} voter, $\beta_{\beta=\mu}^*(\theta_1^{med})$, is to the left of the ideal policy on this diagonal for a θ_2 voter, $\beta_{\beta=\mu}^*(\theta_2^{med})$. Hence, there is a point between $\beta_{\beta=\mu}^*(\theta_1^{med})$ and $\beta_{\beta=\mu}^*(\theta_2^{med})$ that represents the median ideal point restricted to the diagonal. Let us name it M^* . Notice, however, that it does not correspond to $\beta_{\beta=\mu}^*(\bar{\theta})$, since the ideal point on the diagonal for a $\bar{\theta}$ type voter differs depending on whether s/he is in a small or a large district. The indifference line for the district 1 voter that passes through M^* corresponds to a voter with a lower θ than the district 2 voter whose indifference line passes through M^* . Now, by an argument similar to case 2, it is clear that a majority of district 1 voters have ideal lines to the left of M^* , and a majority of district 2 voters have ideal lines to the right of M^* , but exactly half the voters overall have ideal lines to the right of M^* , half to the left. Because the slope of the indifference lines for voters from district 1 is greater than 1, those indifference lines cut the diagonal from below. Therefore, any move to the right of M^* is opposed by a majority, as is any move down. Therefore, M^* is a majority rule equilibrium. It is easily verified that no other point can be a majority rule equilibrium.

¹⁴ There is also the trivial case where the intersection occurs in the $\beta < 0$ region, in which case the majority rule equilibrium is $\beta = \mu = 0$.

REFERENCES

- Alesina, Alberto, and Howard Rosenthal. 1995. *Partisan Politics, Divided Government, and the Economy*. Cambridge, UK: Cambridge University Press.
- Banks, Jeffrey S., and John Duggan. 1998. "Stationary Equilibria in a Bargaining Model of Social Choice." Working Paper. California Institute of Technology.
- Baron, David P., and John Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83(December):1181–206.
- Coughlin, Peter. 1990. *Probabilistic Voting Models*. Cambridge, UK: Cambridge University Press.
- Cr  mer, Jacques, and Thomas R. Palfrey. 1996. "In or Out?: Centralization by Majority Vote." *European Economic Review* 40(1):43–60.

- Grofman, Bernard, and Donald Wittman. 1989. *The Federalist Papers and the New Institutionalism*. New York: Agathon.
- Hinich, Melvin. 1978. "The Mean vs. the Median in Spatial Voting Games." In *Game Theory and Political Science*, ed. Peter Ordeshook. New York: New York University Press. Pp. 357–74.
- Madison, James, Alexander Hamilton, and John Jay. [1788] 1937. *The Federalist Papers*. Modern Library Edition. New York: Random House.
- Ledyard, John. 1984. "The Pure Theory of Large Two-Candidate Elections." *Public Choice* 44(1):7–41.
- Merlo, Antonio. 1997. "Bargaining over Governments in a Stochastic Environment." *Journal of Political Economy* 105(1):101–31.
- Riker, William. 1984. "The Heresthetics of Constitution-Making: The Presidency in 1787, with Comments on Determinism and Rational Choice." *American Political Science Review* 78(March):1–16.
- Riker, William. 1986. *The Art of Political Manipulation*. New Haven, CT: Yale University Press.
- Tsebelis, George, and Jeannette Money. 1997. *Bicameralism*. Cambridge, UK: Cambridge University Press.